Psychological Science

How Many Variables Can Humans Process?
Graeme S. Halford, Rosemary Baker, Julie E. McCredden and John D. Bain

The online version of this article can be found at:

Published by:

$SAGE

http://www.sagepublications.com

On behalf of:

aps
ASSOCIATION FOR
PSYCHOLOGICAL SCIENCE

Association for Psychological Science

Additional services and information for *Psychological Science* can be found at:

Email Alerts: http://pss.sagepub.com/cgi/alerts

Subscriptions: http://pss.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav

>> Version of Record - Jan 1, 2005

What is This?

Research Article

How Many Variables Can Humans Process?

Graeme S. Halford,[1] Rosemary Baker,[1] Julie E. McCredden,[1] and John D. Bain[2]

[1]*University of Queensland, Brisbane, Australia, and* [2]*Griffith University, Brisbane, Australia*

ABSTRACT—*The conceptual complexity of problems was manipulated to probe the limits of human information processing capacity. Participants were asked to interpret graphically displayed statistical interactions. In such problems, all independent variables need to be considered together, so that decomposition into smaller subtasks is constrained, and thus the order of the interaction directly determines conceptual complexity. As the order of the interaction increases, the number of variables increases. Results showed a significant decline in accuracy and speed of solution from three-way to four-way interactions. Furthermore, performance on a five-way interaction was at chance level. These findings suggest that a structure defined on four variables is at the limit of human processing capacity.*

The problem of how to quantify human information processing capacity has been considered crucial at least since the article by Miller (1956). Optimal learning depends on reducing the complexity of information to a level that does not exceed capacity (Elman, 1993). Reasoning tasks must also be coded by the problem solver so that no step in the solution exceeds processing capacity (Birney & Halford, 2002). Furthermore, the number of variables that can be integrated into a single cognitive representation is a major constraint on cognitive and neuropsychological processes.

The limits to processing have been estimated theoretically (Christoff et al., 2001; Halford, Wilson, & Phillips, 1998; Hummel & Holyoak, 2003; Phillips & Niki, 2002). However, although there are capacity estimates for visual working memory and short-term memory (Cowan, 2001; Luck & Vogel, 1997), there has not been a successful empirical determination of processing capacity for variables.

Assessment of processing capacity is difficult because of the great power of strategies for reducing processing load, thereby optimizing use of available capacity (Hirst, Spelke, Reaves, Caharack, & Neisser, 1980). Therefore, problem-solving strategies that reduce complexity, though of immense value in other contexts, must be constrained as far as possible in order to determine underlying capacity.

Capacity to process variables can be assessed by requiring participants to interpret certain types of graphical representations of interactions among variables. In an interaction, the effect of any variable is modified by the effects of all the other variables. Interpretation of an interaction thus requires that information derived from all variables be integrated into a single complex concept. Higher-order interactions (i.e., those with more than two independent variables) create high processing loads because, although some serial processing might occur, accurate interpretation cannot be based on a subset of variables.

Principles embodied in Bertin's (1977/1981) theory of graph perception and in theories of graph processing (Carswell, 1992; Gillan & Lewis, 1994; Pinker, 1990) are normally used to facilitate interpretation of graphs. However, graphs can be designed to preclude reductions in processing load, as is the case when bar graphs describe interactions. In such graphs, the tops of the bars cannot easily be described by a single predicate such as "concave increasing" (Pinker, 1990). That is, they cannot be chunked into a unit describing an overall trend or configuration. Thus, using bar graphs to represent interactions helps to constrain strategies that would reduce the number of variables being processed.

For example, to understand the effect of the variable *chocolate versus carrot* on the preference for cakes within each half of Figure 1a, one must also understand the effect of the variable *fresh versus frozen*. In order to map the descriptive sentences in the Figure 1a caption onto the bar graphs, one must group the bars into structures so that their relative heights relate the variables to one another in the same way as the variables are related conceptually by the sentences. In the current study, we

Address correspondence to Graeme Halford, School of Psychology, University of Queensland, Brisbane, Queensland 4072, Australia; e-mail: gsh@psy.uq.edu.au.
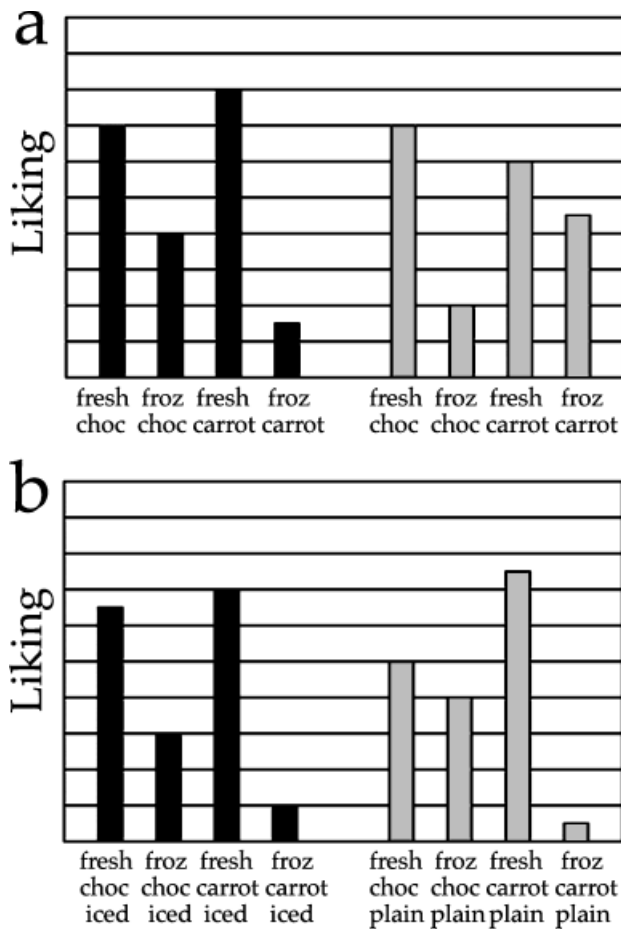
Fig. 1. Examples of problems with eight bars. For each problem, participants viewed a bar graph and an accompanying verbal description. (The graphs were presented in blue and yellow, represented here by black and gray, respectively.) The task was to indicate whether "greater" or "smaller" would correctly complete the final sentence of the description. Horizontal lines represent 20 units. In the 2 × 2-way problem shown here (a), the verbal description presented to participants was as follows: "People prefer **fresh** cakes to **frozen** cakes. The difference depends on the flavor (**chocolate** vs **carrot**). Left half(blue): The difference between **fresh** and **frozen** is (greater/smaller) for **chocolate** cakes than for **carrot** cakes. Right half(yellow): The difference between **fresh** and **frozen** is (greater/smaller) for **chocolate** cakes than for **carrot** cakes." (The correct answers for this problem are "smaller" and "greater," respectively.) In the 3-way problem shown here (b), the verbal description presented to participants was as follows: "People prefer **fresh** cakes to **frozen** cakes. The difference depends on the flavor (**chocolate** vs **carrot**) and the type (**iced** vs **plain**). The difference between **fresh** and **frozen** increases from **chocolate** cakes to **carrot** cakes. This increase is (greater/smaller) for **iced** cakes than for **plain** cakes." (The correct answer is "smaller.")

used these types of sentence-graph mappings to determine processing ability for increasing levels of complexity.

We aimed to use the noncondensable characteristics of bar-graph representations to investigate the limits to human processing capacity. Our complexity estimates of the graphical representations for the interactions were based on the relational complexity metric, which has been applied to a wide range of cognitive domains (Andrews & Halford, 2002; Birney & Hal-

ford, 2002; Christoff et al., 2001; Halford et al., 1998; Kroger et al., 2002; Waltz et al., 1999). In this metric, cognitive complexity is defined by the *arity* (i.e., number of arguments, or slots) of the relations that are represented by the participant in order to perform the task. An *n*-ary relation is a set of points in *n*-dimensional space, so *arity* of a concept corresponds to the number of dimensions encompassed by the relation. Relational complexity theory proposes that the amount of information that has to be processed in a single cognitive step can be reduced by *conceptual chunking* into fewer, larger entities or by *segmentation* into smaller subtasks that are performed serially (Halford et al., 1998; Simon, 1974). Both conceptual chunking and segmentation are constrained in interactions because of the need to process the variables jointly.

From a mathematical perspective, an *n*-way interaction corresponds to a set of values of the dependent variable, which is a function of the *n* independent variables defining the interaction. From a cognitive perspective, graphically represented interactions can be represented as the effect on the different levels of one variable due to the influence of other variables, that is, how the effect of variable *A* is modified by the effect of variable *B*, and then how this modification is further affected by variable *C*, and so forth. For example, for a problem that involves only binary variables, if the effect due to the two levels of variable *A* is called A_{diff}, then in a two-way interaction, A_{diff} is moderated by variable *B*, an effect that is represented by $B(A_{\text{diff}})$. In the three-way interaction, $B(A_{\text{diff}})$ is moderated by variable *C*, an effect that is represented by $C(B(A_{\text{diff}}))$, and so on. In this formulation, *A*, *B*, *C*, and so forth are operator variables, and each operator creates two difference sizes for its argument, as follows:

$$A_{\text{diff}} \rightarrow A_1, A_2 \text{ (two bar heights on a graph)}$$
$$B(A_{\text{diff}}) \rightarrow A_{\text{diff1}}, A_{\text{diff2}} \text{ (four bar heights on a graph)}$$
$$C(B(A_{\text{diff}})) \rightarrow C(A_{\text{diff1}}, A_{\text{diff2}}) \rightarrow A_{\text{diff11}}, A_{\text{diff12}}, A_{\text{diff21}}, A_{\text{diff22}}$$
$$\text{(eight bar heights on the graph), and so on.}$$

To understand an interaction, the problem solver must understand all the points on the graph not as individual points, but as collections that define a cognitive structure. In order to minimize the cognitive load imposed by this structure, the problem solver may make use of operator variables to generate all values that are part of the interaction. In this case, the two two-way graphs in Figure 1a are represented not as separate bar heights, but as sets of pairs of heights. Within each pair, the height difference between *fresh* and *frozen* is represented as a single entity, such as preference—that is, in Pinker's (1990) terms, as a single predicate. This preference is then influenced by the variable *chocolate versus carrot*. It is further influenced by the variable *iced versus plain* in the three-way graph (Fig. 1b), and even further influenced by the variable *rich versus low fat* in the four-way graph (Fig. 2b).
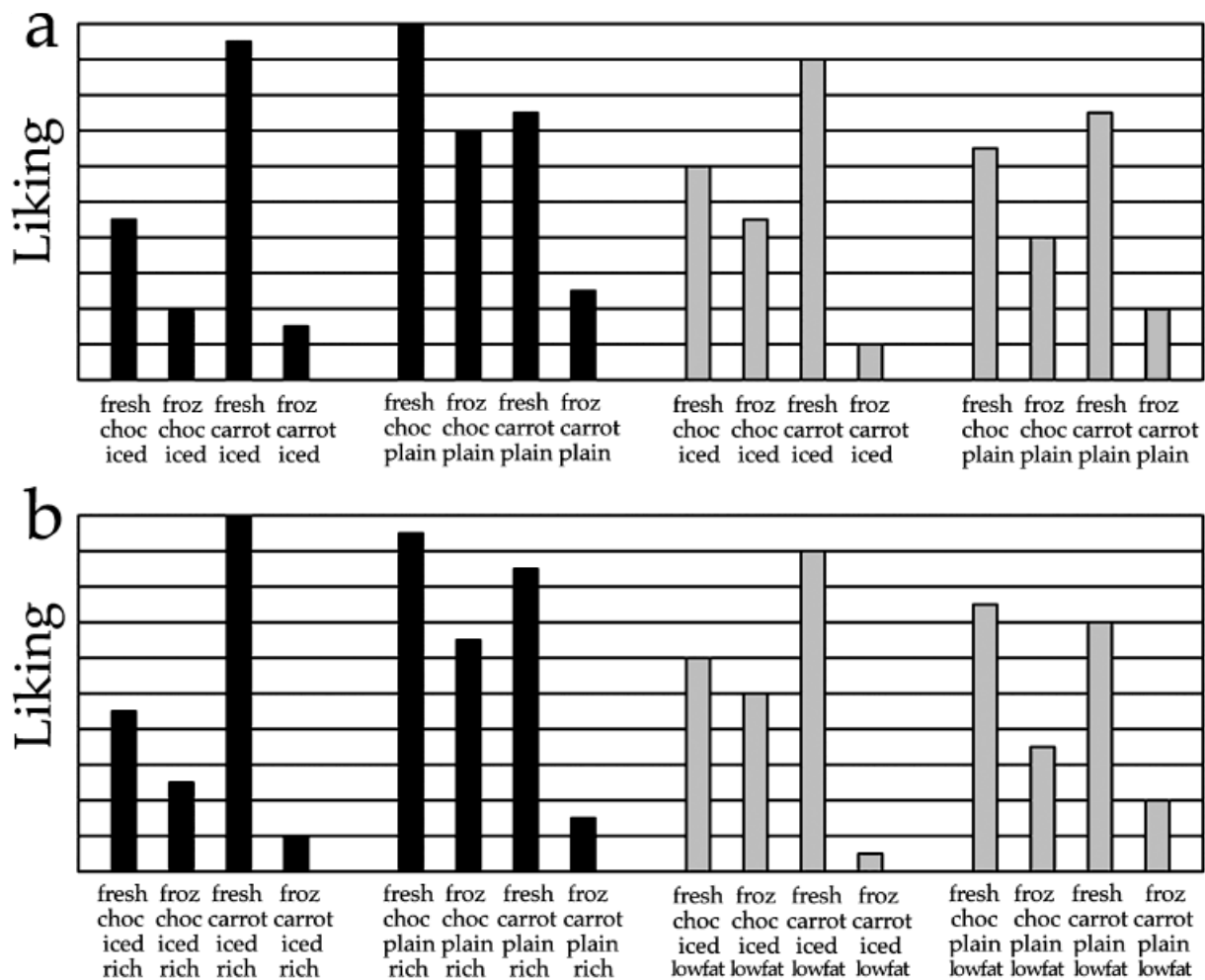
Fig. 2. Examples of problems with 16 bars. For each problem, participants viewed a bar graph and an accompanying verbal description. (The graphs were presented in blue and yellow, represented here by black and gray, respectively.) The task was to indicate whether "greater" or "smaller" would correctly complete the final sentence of the description. Horizontal lines represent 20 units. In the 2 × 3-way problem shown here (a), the verbal description presented to participants was as follows: "People prefer **fresh** cakes to **frozen** cakes. The difference depends on the flavor (**chocolate** vs **carrot**) and the type (**iced** vs **plain**). The difference between **fresh** and **frozen** increases from **chocolate** cakes to **carrot** cakes. Left half(blue): This increase is (greater/smaller) for **iced** cakes than for **plain** cakes. Right half(yellow): This increase is (greater/smaller) for **iced** cakes than for **plain** cakes." (The correct answers for this problem are "greater" and "greater," respectively). In the 4-way problem shown here (b), the verbal description presented to participants was as follows: "People prefer **fresh** cakes to **frozen** cakes. The difference depends on the flavor (**chocolate** vs **carrot**), the type (**iced** vs **plain**) and the richness (**rich** vs **low fat**). The difference between **fresh** and **frozen** increases from **chocolate** cakes to **carrot** cakes. This increase is greater for **iced** cakes than for **plain** cakes. There is a (greater/smaller) change in the size of the increase for **rich** cakes than for **low fat** cakes." (The correct answer is "smaller.")

Although an operator-based representation removes the need to represent all bar heights in order to understand the interaction, it still requires the simultaneous representation and integration of all the variables on which the interaction is defined. Thus, the full cognitive representation of an *n*-way interaction requires the simultaneous coding of *n* variables. The two-, three-, and four-way interactions described in Figures 1 and 2 thus require two, three, and four variables, respectively, to be related within a single concept.

The complexity of interactions that an individual is able to interpret is thus a direct measure of the number of variables that the person can process at one time. This measure holds even if one elects to represent structures hierarchically, because the complexity of hierarchical structures can be defined by the number of variables required for representation (Halford et al., 1998). In principle, it also holds if the variables are other than binary because, at least with monotonic effects, each variable can be coded by the magnitude of its effect.

In the study described next, we used graphical representations of interactions of varying orders of complexity to manipulate the number of variables that needed to be considered in one problem-solving decision. In these manipulations, the memory load was equalized as far as possible, so that it was processing load rather than storage that was varied. We ex-

pected that as complexity of interactions increased, processing difficulty would increase, as measured by number of errors, solution times, and confidence ratings.

EXPERIMENT 1

The task in this experiment was to interpret graphical presentations of two-, three-, and four-way interactions, which require two, three, and four variables, respectively, to be processed. Specifically, the task required participants to choose whether "greater" or "smaller" would be the correct completion of the final sentence in a verbal description of each interaction.

Method

Participants
To optimize expertise for the task, we recruited 30 participants who were academic staff and graduate students in psychology and computer science and had experience in interpreting the type of data presented.

Materials
Each problem involved selecting the correct form of a verbal description to match a graphical representation of an interaction based on fictitious data on one of six everyday topics (see Figs. 1 and 2). The verbal descriptions were written so as to suggest that the lowest level of difference between pairs be treated as a single entity, such as a preference or difference (e.g., "People prefer **fresh** cakes to **frozen** cakes"). The sentences then described how the other variables in the interaction affected that preference entity (e.g., "The difference between **fresh** and **frozen** increases from **chocolate** cakes to **carrot** cakes"). Thus, the construction of the sentences encouraged conceptual constructions as described in the representational analysis just presented, in which a difference between two levels of one variable is operated on by further variables. To keep the complexity of the tasks equal at all levels of interaction, we used only binary variables.

The materials were designed to equalize all task characteristics except for complexity. In the crucial comparisons, the input memory load was kept as constant as possible by equalizing the amount of verbal information and the number of bars. Thus, 2 two-way interactions were compared with a three-way interaction using 8 bars (as shown in Figs. 1a and 1b) and 2 three-way interactions were compared with a four-way interaction using 16 bars (as shown in Figs. 2a and 2b).

The general configuration of bars was consistent across problems, regardless of level of complexity. In the 2×2-way and 2×3-way displays, two graphs were presented side by side, one in blue and one in yellow (shown as black and light gray, respectively, in Figs. 1 and 2). In the single 3-way and 4-way displays, the left half of the bars were shown in blue and

the right half in yellow to ensure comparability with the problems involving pairs of graphs.

There was no repetition of particular sets of height values across graphs, to ensure that all problems were independent of one another. However, the graphs were designed to be as consistent as possible in structure, both to eliminate extraneous variables and to encourage optimal performance. For example, the lower-level interactions within a higher-level interaction were consistent in direction, differing only in magnitude. This enabled participants to hold constant the directions of the nested differences and to focus on their relative sizes to determine the direction of the highest-level interaction (and hence choose the solution "greater" or "smaller"). This consistency in the direction of differences was maintained across all graphs, so that problems of different levels of complexity differed essentially only in the number of factors.

To ensure that height differences could be discerned visually, in all graphs we maintained minimum differences between pairs of bars representing preferences. For example, if one assigns an arbitrary value of 20 units to each interval between horizontal lines in Figures 1 and 2, then the minimum difference between pairs was 20. Sizes of interactions were also designed to aid discernibility. The magnitude of each interaction was defined by the size difference at the highest level of the interaction. For example, for the left (black) two-way graph in Figure 1a, the highest-level difference is $(140-80)-(160-30) = -70$ units. Across all problem types, three size differences, corresponding to 50, 60, and 70 units, were used at the highest level of interaction (e.g., all the sample graphs in Figs. 1 and 2 represent an absolute size difference of 70 at the highest level). Furthermore, minimum differences were maintained for the lower-level interactions. The minimum difference for two-way interactions nested within the three-way and four-way interaction problems was 20. The minimum difference for three-way interactions nested within four-way interaction problems was 50.

Each participant was assigned two of the three difference sizes at the highest level of interaction (50, 60, 70), each occurring equally often over the set of experimental problems. Each person was assigned just one topic for all problems administered, to minimize the amount of new information to be assimilated for each new problem.

The materials were constructed to reduce the likelihood that simplification strategies such as elimination or chunking could be used. Within problems, no bar height was repeated within sets of bars requiring joint processing, and there was no repeated magnitude of difference at any given level of interaction. Horizontal grid lines aided comparison of bar heights, but there was no numbering, and calculation was explicitly discouraged.

Procedure
All problems were presented on a laptop computer. The verbal description of an interaction was presented first, one statement to a line, to facilitate reference. This description contained two

possible directions of the interaction. A graph was then displayed beneath the text, initially using equal bar heights, to permit mapping of the descriptor variables onto the corresponding bar labels. To display the bar heights corresponding to the values for the variables (as in Figs. 1 and 2), participants held down two widely spaced keys with the first finger of each hand. This precluded recoding of heights or differences using fingers. Then participants released the keys (causing the bar heights to become equal again) and used the mouse to select either "greater" or "smaller" as the correct completion of the last sentence of the verbal description for that problem. The verbal description remained on the screen throughout, so that participants did not need to retain additional data in working memory while solving the problem. Solution times were recorded for the problem-solving phase (i.e., while the bars were unequal length). Participants then rated their confidence in each answer on a scale from 0 (*pure guess*) to 5 (*fully confident*).

Problems were administered in a single session that comprised (a) one demonstration 2 × 2-way problem; (b) one practice 2 × 2-way problem; (c) one demonstration 3-way problem; (d) one practice 3-way problem; (e) a set of four experimental problems, consisting of two 2 × 2-way problems and two 3-way problems, presented in random order; (f) one demonstration 2 × 3-way problem; (g) one practice 2 × 3-way problem; (h) one demonstration 4-way problem; (i) one practice 4-way problem; (j) a set of four experimental problems, consisting of two 2 × 3-way problems and two 4-way problems, presented in random order; and (k) re-presentation of the last 4-way problem and 2 × 3-way problem, with submitted answers shown, for tape-recording of a verbal protocol.

The demonstration and practice phases were designed to provide adequate training and familiarity with the structure of the problems, while avoiding mental fatigue or the acquisition of strategies that would circumvent the requirements of the task. Therefore, problems of the different types were introduced systematically, building up by one factor at a time from two-way to three-way, and then from three-way to four-way, using the same topic as for the experimental problems. The description, graph, and chosen answer reappeared on the screen after completion of each demonstration and practice problem, to provide feedback. No feedback was provided for experimental problems. Participants were invited to rephrase problems in their own words if they wished and were offered paper and pencil to do so, but this option was never used.

Results

Table 1 shows that as the order of the interaction increased, the number of participants answering one or more problems incorrectly increased. For 2 × 3-way and 4-way problems, the McNemar change test showed that this pattern was significant, $\chi^2(1, N = 30) = 6.25$, $p < .02$.

TABLE 1

Number of Participants Answering Neither, One, or Both Problems of Each Type Correctly

	Problem type			
Score	2 × 2-way	3-way	2 × 3-way	4-way
Both correct	30	26	23	13
One incorrect	0	4	7	13
Both incorrect	0	0	0	4

Note. $N = 30$.

Mean solution times and confidence ratings are given in Figure 3, which shows that as task complexity increased, speed and confidence both decreased, particularly between the 2 × 3-way and 4-way problems. For solution times, one-tailed t tests revealed that for the 8-bar graphs, there was no significant difference between the two levels of complexity (2 × 2-way vs. 3-way), $t(29) = 0.23$, n.s., $d = 0.22$, whereas for the 16-bar graphs, there was a significant difference between the two levels of complexity (2 × 3-way vs. 4-way), $t(29) = 3.83$, $p < .05$,
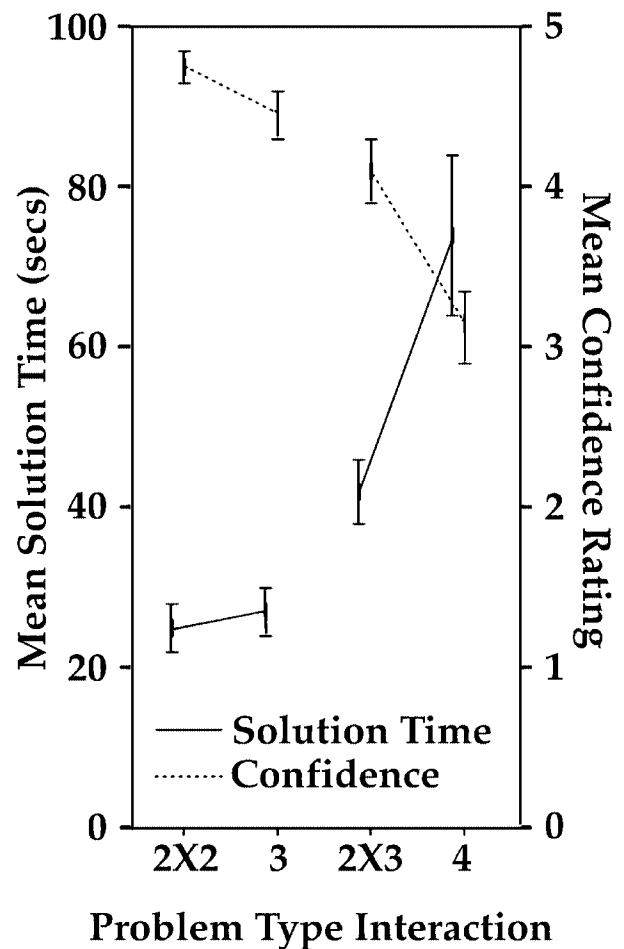

Fig. 3. Mean solution times and confidence ratings as a function of problem type in Experiment 1.

$d = 0.68$. For confidence ratings, there were significant differences between the 2 × 2-way and the 3-way problems, $t(29) = 4.04$, $p < .05$, $d = 0.71$, and between the 2 × 3-way and the 4-way problems, $t(29) = 5.04$, $p < .05$, $d = 0.89$. Variations in the graphical form of the interaction used did not influence the findings: The three sizes of the difference at the highest level did not interact with the observed effects.

A curve fit applied to the percentage of correct responses yielded the quadratic function $P = 51 + 45.5N - 10.5N^2$ (where P = percentage correct and N = level of interaction). This function accounts for all of the variance ($R^2 = 1.00$). The function extrapolates to chance level (50%) between four-way and five-way problems. This prediction was tested in Experiment 2.

The mean solution times and confidence ratings calculated only from problems answered correctly were almost identical to those shown in Figure 3 for the complete data set. The largest differences were for the four-way problems, for which the mean solution time for correct answers was 78.28 s, as opposed to 76.85 for all problems, and the mean confidence rating was 3.25 instead of 3.12.

The verbal protocols for the three- and four-way interactions indicated that processing-load difficulties, reported by 10 of the 30 participants, were exclusive to four-way problems (e.g., "This is what I'm having trouble holding onto," "Everything fell apart and I had to go back," "I kept losing information").

EXPERIMENT 2

The task in this experiment was to interpret a graphical representation of a five-way interaction, which required five variables to be processed.

Method

Participants
Twenty-two of the 30 participants from Experiment 1 were available to take part in Experiment 2. In terms of their performance on the problems, Experiment 2 participants were representative of the original sample: Ten had previously answered both four-way problems correctly, 10 had answered one correctly, and 2 had answered both incorrectly (out of 13, 13, and 4, respectively, for the original sample).

Materials and Procedure
Participants were presented with a five-way problem, constructed using 2 four-way bar graphs, one blue and one yellow, presented side by side on paper, with labels and a full written description of both interactions. Each person received a problem involving the same topic as in the first experiment. Participants were asked to say whether the interaction was larger in the blue or the yellow graph, which were said to represent different (fictional) surveys. Thus, the problem was similar to those in the previous experiment, but with a fifth factor added.

Results
Twelve participants gave correct responses, and 10 gave incorrect responses, which is no better than chance. This result is consistent with the extrapolation from Experiment 1. The mean confidence rating was 2.28, which is lower than the mean confidence rating for the four-way interactions. Seven of the 22 participants gave explicit verbal indications of processing-load difficulties (e.g., "It becomes too much"; "When I got to the 5th level, I just lost track"; "My brain wouldn't really do the comparison at that level").

DISCUSSION

All dependent measures, that is, the number of correct problems, solution times, and confidence ratings, indicated a disproportionate decline in performance from 2 × 3-way interactions to 4-way interactions, as compared with the transition from 2 × 2-way interactions to 3-way interactions. Furthermore, verbal indications of processing-load difficulties were exclusive to 4-way problems and 5-way problems. Given the controls that were built into the design of the graphs, which minimized differences due to storage or configurations, we can conclude that the increased level of difficulty for the 4- and 5-way problems was due to increased processing loads.

The results show that a four-way interaction is difficult even for experienced adults to process without external aids. A soft limit is indicated by the decline from three- to four-way interactions, with five-way problems being performed no better than chance. A four-way interaction requires four variables to be integrated in a representation, so our findings suggest that a structure defined on four variables is at the limit of human processing capacity. This limit of four variables would coordinate well with visual and short-term memory capacities of four items (Cowan, 2001; Luck & Vogel, 1997) and is consistent with predictions from symbolic connectionist models (Halford et al., 1998; Hummel & Holyoak, 2003).

Processing loads required for the 2 × 3-way and 4-way problems differed because two 3-way problems can be processed independently, and a solution can be stored for each, whereas the two halves of a 4-way problem must be processed relative to each other, and cannot be decomposed into separate problems. Therefore, the increase in working memory load from the 2 × 3-way to the 4-way problems was not simply due to the amount of information that was stored, but was due to the number of variables that had to be related in the representations of the problems.

The results imply that the strategies for reasoning and decision making must entail processing of no more than four variables in any one cognitive step. However, more complex tasks can be processed by a number of well-established means. Conceptual chunking (Andrews & Halford, 2002; Halford et al., 1998) is analogous to collapsing over variables in analysis of variance and can be used to reduce the number of variables

processed in one step, but at the cost that relations between chunked variables become temporarily inaccessible. For example, velocity equals distance divided by time ($V = dt^{-1}$). However, if velocity is reduced to a single variable, as in the scale on a speedometer, one cannot answer questions such as "What would happen to an object's velocity if the object travels the same distance in half the time?" It is possible to answer such questions only by processing three variables, according to the formula $V = dt^{-1}$. One can chunk velocity as a single variable in order to process acceleration as a difference between velocities at two times, t_1 and t_2 ($A = V_1 - V_2$). Then acceleration can be chunked as a single variable to define force as mass times acceleration ($F = MA$), and so on. Alternatively, one can segment complex tasks into simpler subtasks that are performed serially, but representations in any one segment will not include relations to variables in other segments. Therefore, effective problem-solving strategies must include a sequence of steps that process all the relevant relations between variables, but never require more than four variables in one step.

It is a major function of expertise to recognize higher-order variables that relate chunked representations of lower-order variables. Velocity is a higher-order variable that relates time and distance, acceleration relates changes in velocity to time, and force relates acceleration and mass. Variance, defined as the mean squared deviations from a mean, is an example of a higher-order variable that forms part of the expertise of psychological scientists. Knowledge of higher-order variables can be applied to interpretation of interactions, by chunking values of two or more variables into a trend or a configuration. However, our findings suggest that the underlying cognitive processes in such tasks represent a maximum of approximately four variables.

Higher-order variables are used to overcome capacity limitations in both academic and applied contexts, especially in complex tasks such as air-traffic control (Boag, 2003). The tasks that are most intractably difficult are those in which conceptual chunking and segmentation are constrained, as in interpretation of interactions, knights-and-knaves tasks (Birney & Halford, 2002), and some problems that have been prominent in cognitive development research (Andrews & Halford, 2002).

REFERENCES

Andrews, G., & Halford, G.S. (2002). A cognitive complexity metric applied to cognitive development. *Cognitive Psychology, 45*, 153–219.

Bertin, J. (1981). *Graphics and graphic information processing* (W.J. Berg & P. Scott, Trans.). Berlin, Germany: Walter de Gruyter. (Original work published 1977)

Birney, D.P., & Halford, G.S. (2002). Cognitive complexity of suppositional reasoning: An application of the relational complexity metric to the knight-knave task. *Thinking & Reasoning, 8*, 109–134.

Boag, C. (2003). *Investigating the cognitive determinants of expert performance in air traffic control*. Unpublished doctoral dissertation, University of Queensland, Brisbane, Australia.

Carswell, C.M. (1992). Reading graphs: Interactions of processing requirements and stimulus structure. In B. Burns (Ed.), *Percepts, concepts and categories: The representation and processing of information* (pp. 605–645). Amsterdam: North-Holland.

Christoff, K., Prabhakaran, V., Dorfman, J., Zhao, Z., Kroger, J.K., Holyoak, K.J., & Gabrieli, J.D.E. (2001). Rostrolateral prefrontal cortex involvement in relational integration during reasoning. *NeuroImage, 14*, 1136–1149.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*, 87–185.

Elman, J.L. (1993). Learning and development in neural networks: The importance of starting small. *Cognition, 48*, 71–79.

Gillan, D.J., & Lewis, R. (1994). A componential model of human interaction with graphs: 1. Linear regression modeling. *Human Factors, 36*, 419–440.

Halford, G.S., Wilson, W.H., & Phillips, S. (1998). Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. *Behavioral and Brain Sciences, 21*, 803–831.

Hirst, W., Spelke, E.S., Reaves, C.C., Caharack, G., & Neisser, U. (1980). Dividing attention without alternation or automaticity. *Journal of Experimental Psychology: General, 109*, 98–117.

Hummel, J.E., & Holyoak, K.J. (2003). A symbolic-connectionist theory of relational inference and generalization. *Psychological Review, 110*, 220–264.

Kroger, J., Sabb, F.W., Fales, C., Bookheimer, S.Y., Cohen, M.S., & Holyoak, K. (2002). Recruitment of anterior dorsolateral prefrontal cortex in human reasoning: A parametric study of relational complexity. *Cerebral Cortex, 12*, 477–485.

Luck, S.J., & Vogel, E.K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390*, 279–281.

Miller, G.A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*, 81–97.

Phillips, S., & Niki, K. (2002). Separating relational from item load effects in paired recognition: Right temporo-parietal and middle frontal gyral activity with increased associates, but not items during encoding. *NeuroImage, 17*, 1031–1055.

Pinker, S. (1990). A theory of graph comprehension. In R. Freedle (Ed.), *Artificial intelligence and the future of testing* (pp. 73–126). Hillsdale, NJ: Erlbaum.

Simon, H.A. (1974). How big is a chunk? *Science, 183*, 482–488.

Waltz, J.A., Knowlton, B.J., Holyoak, K.J., Boone, K.B., Mishkin, F.S., de Menezes Santos, M., Thomas, C.R., & Miller, B.L. (1999). A system for relational reasoning in human prefrontal cortex. *Psychological Science, 10*, 119–125.